

The Paradox of Exploitation

Benjamin Ferguson¹

Received: 6 January 2015 / Accepted: 23 September 2015 / Published online: 8 October 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The concept of exploitation brings many of our ordinary moral intuitions into conflict. Exploitation—or to use the commonly accepted ordinary language definition, taking unfair advantage—is often thought to be morally impermissible. In order to be permissible, transactions must not be unfair. The claim that engaging in mutually beneficial transactions is morally better than not transacting is also quite compelling. However, when combined with the claim that morally permissible transactions are better than impermissible transactions, these three imply the counterintuitive claim that it is obligatory to engage in mutually beneficial transactions. In this paper I outline the conditions that comprise this ‘paradox of exploitation’ along with a solution that involves replacing one of the original conditions with a condition I call Weak Non-worseness. The solution captures the priority of our concerns about exploitation by making a concern for the fairness of a transaction subsidiary to a concern for the welfare of the would-be exploited.

Recently the concept of exploitation has received renewed attention. Much of the philosophical discussion has centred on debates about what makes a transaction fair.¹ This is an important question since exploitation is generally thought to involve unfairness in transaction. Most of the remaining debate surrounds questions about whether conditions in addition to fairness—such as ‘attitudinal deficiencies’ on the

¹ See, for example, Steiner (1984), Wertheimer (1996), Vrousalis (2013), Goodin (1987), and Roemer (1982).

✉ Benjamin Ferguson
b.r.ferguson@vu.nl

¹ Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

part of the exploiter—are also necessary for exploitation.² Both of these questions are concerned primarily with identifying the scope of exploitation, with *what exploitation is*. However, as G.A. Cohen writes, we need more than “a precise definition of what exploitation is...we want to know exactly why it is wrong.”³ Not only has this latter task been largely neglected, but as I will show, providing an account of *why* exploitation is wrong is also much more difficult than it might first appear.

Attempts to identify what is wrong with exploitation are often frustrated because the concept brings many of our deeply held moral intuitions into conflict. Because they incorporate intuitions about the impermissibility of engaging in unfair transactions alongside a desire to improve the welfare of those involved in the transactions, accounts of exploitation’s wrongness include concerns typically attributed to both non-consequentialist and consequentialist moral frameworks. This combination makes it difficult to articulate which moral principles are violated in cases of exploitation.

In this article, I describe a paradox that arises when we attempt to account for four normative intuitions about transactions that are Pareto improving and consensual, but nevertheless unfair. The paper proceeds as follows: in the Sect. 1 I outline a basic tension between a concern for fairness and a concern for welfare in exploitative transactions. In the Sect. 2 I show that an attempt to incorporate both of these concerns leads to what I call the ‘paradox of exploitation’. In Sect. 3 I evaluate the four conditions that generate the paradox and outline possible resolutions. In Sect. 4 I advocate a resolution that, I argue, offers the best balance between a concern for fairness and a concern for welfare. Finally, the Sect. 5 concludes.

1 The Initial Tension

When someone steals my wallet, they take it without consent and make me worse off than I was. Conversely, when I buy an apple from a market stall, the transaction is (ordinarily) consensual and mutually beneficial. Some accounts of exploitation include the first kind of interaction in the scope of exploitation. They allow exploitation to extend to non-consensual or harmful transactions. Other accounts limit their scope to consensual and mutually advantageous transactions.⁴ I do not wish to take a stand on whether these broader forms of interaction are ‘really’ exploitations. The focus of this paper is on transactions commonly described as exploitative, but which are both consensual and mutually beneficial. More than harmful or non-consensual transactions, mutually beneficial and consensual transactions present a puzzle, partly because the following claim has some *prima facie* appeal:

² See, for example, Sample (2003), Goodin (1987), Vrousalis (2013), and Veneziani (2013).

³ Cohen (1995, 144).

⁴ Steiner (1984, 1987) restricts the scope of exploitation to mutually beneficial transactions. See Wertheimer (2007) and (1996, 14) for discussions of whether ‘exploitation’ should be applied only to mutually beneficial transactions.

Pareto Permissibility If a transaction is consensual and strictly Pareto improving, then it is morally permissible.⁵

In consensual and strictly Pareto improving transactions, both parties are better off having transacted and no other parties are burdened with negative externalities.⁶ Further, when both parties token consent to the transaction, concerns about paternalism and certain forms of force do not arise.⁷ In what follows I will refer to Pareto improving, consensual, but nevertheless unfair transactions as exploitations.⁸

The Pareto Permissibility claim implies that some paradigmatic cases of exploitation are permissible. Consider a sweatshop labour contract. Here, perhaps more than anywhere else, the term ‘exploitation’ is used to condemn a practice that many argue is contemptible and morally amiss. Sweatshop labourers work long hours at near-subsistence wages in order to produce household goods that the workers cannot themselves afford to buy. Robert Ross’s account of conditions at Chentex, a Nicaraguan garment firm subcontracted by retail chain stores is, sadly, typical of this form of labour:

Chentex workers earn less than 1 percent of the retail price of the jeans they stitch—between thirty and forty cents an hour. This compares to the 10 percent typical of the global north.

When we visit the workers’ homes...we meet a woman, Christina, who was fired from Chentex...Christina’s home is a wooden frame, ten-feet square, hung with plastic sheeting for two of the walls and with cardboard boxes that once held shirts shipped from the freezone in Panama for the rest. Her shack has a dirt floor and holds one large bed and (barely) two chairs for herself, her husband, and their baby. Her toilet is a hole in the ground.⁹

Christina’s condition is appalling and her wages—when she was employed—were low (around \$4 per day), but her story is tragically common.

Nevertheless, workers in Chentex factories often receive pay that exceeds wages offered elsewhere in Nicaragua. Ross notes that per capita income in Nicaragua “is

⁵ It is important to distinguish moral (im)permissibility from legal prohibition. Acts may be morally impermissible—in the sense that there are moral reasons persons ought to refrain from performing them—without being legally prohibited. For example, adultery and lying may be wrong, but it is often presumed that there are good reasons for the state to refrain from legally prohibiting these acts. In this paper I am interested solely in the question of *moral* permissibility. This distinction between moral permissibility and legal permissibility is discussed in Wertheimer (1996, chapter 9). Wertheimer calls questions about moral permissibility questions about ‘moral weight’ and questions about the legal prohibition of acts questions about ‘moral force’.

⁶ Throughout this paper I will always assume the strict version of Pareto improvement.

⁷ We can imagine a Pareto improving transaction to which Bob would have consented, but did not *actually* consent, perhaps because Carol made the transaction on Bob’s behalf (and without his permission) or because Alice forced Bob to transact. In both cases it seems that the lack of actual consent gives Bob a valid complaint against Carol (and Alice). However, note that for revealed preference theorists and subjective welfarists, strict Pareto implies hypothetical consent.

⁸ This definition is purely conventional—as noted above, I do not want to take a stand here on issues of scope.

⁹ Ross (2004, 116–18).

about \$470 annually,” but garment workers can earn two or three times as much.¹⁰ Consequently, jobs like Christina’s are in high demand. Not only do Chentex workers earn more than their compatriots, they also freely choose—in fact, compete—to work in sweatshop conditions. Christina’s employment at Chentex was both consensual and Pareto improving. Thus, according to Pareto Permissibility the transaction was morally permissible. Yet, many would resist this conclusion. Instead, they would draw the opposite conclusion, outlined in the following claim:

Moral Impermissibility There exist some consensual and strictly Pareto improving transactions (e.g., sweatshop contracts) that are morally impermissible.

Pareto Permissibility and Moral Impermissibility obviously conflict. The latter is simply the negation of the former. It cannot be the case that all strictly Pareto improving and consensual transactions are permissible and that the transaction between Christina and Chentex is strictly Pareto improving, consensual, but nevertheless, morally impermissible.

One response to the tension between Pareto Permissibility and Moral Impermissibility, which I will address briefly before setting aside, is to claim that there are no transactions that actually satisfy Moral Impermissibility; either because the consent of those in positions of severe disadvantage cannot be considered *true* consent, or because although they appear beneficial, exploitative transactions are not *really* Pareto improving. These claims are certainly true of some, and perhaps many cases. However, it is implausible that all sweatshop contracts ordinarily described as exploitative are, all things considered, harmful. A similar response can be given to those who deny that workers’ consent is *true* consent when their choice conditions are constrained. In addition to being paternalistic (by denying Christina’s own claims that she is better off working at Chentex), these responses are rather unimaginative. Even if no *actual* cases of sweatshop labour are consensual and strictly Pareto improving, it seems possible to imagine transactions that are.

Note that this dismissal comes with two important qualifications. The first is that I do not claim that determining whether Pareto and consent are satisfied in certain cases is unimportant. It may be that nearly all cases ordinarily called exploitations are either non-consensual or harmful and thus wrong for precisely these reasons. Many labour contracts that initially appear beneficial, may turn out to be harmful to the transactors, or to involve negative externalities, harming others. In these cases the puzzle of how the transactions can be morally impermissible is solved by appealing to their harm. However, my primary interest in this paper is with the question of whether (and how) transactions that are Pareto improving and consensual can be impermissible.

The second is that I do not deny the consent given in cases of exploitation is, in some way, morally suspect. I shall assume that those involved in transactions we call exploitations *token* consent and that this consent is suitably informed and competent: these individuals know the relevant facts about the transaction in which

¹⁰ Ross (2004, 114).

they are about to engage.¹¹ However, many will object that this tokened consent is invalid or not morally transformative because it fails to satisfy certain voluntariness conditions. Yet the violation of these voluntariness conditions are also those responsible for the intuition that some transactions are unfair. For example, if we say that Christina is unjustly disadvantaged vis à vis Chentex we may think that though she tokens (informed) consent to her job, her consent is less-than-ideal. A non-exploitative transaction would be one in which she was not disadvantaged in such a way that her consent was compromised. And this would lead to different terms than those she faces in the actual transaction. Thus, we can say that the terms of the actual transaction are *unfair* compared to the counterfactual fair transaction. Consequently, I think these concerns about the validity of consent in the sense of morally transformative consent fall under the category of unfairness.

With these qualifications in place, consider the tension between Moral Impermissibility and Pareto Permissibility. A handful of commentators have (somewhat notoriously) dismissed Moral Impermissibility, concluding that mutually beneficial and consensual transactions simply are not impermissible. A startling example is Nicholas Kristof's claim that "the central challenge in the poorest countries is not that sweatshops exploit too many people, but that they don't exploit enough."¹² A similar position is defended by Matt Zwolinski, who claims that "since ...providing no monetary benefit does not violate anyone's rights, and since a contract whereby sweatshops agree to provide some benefit does not in itself violate anyone's rights, it follows that such contracts are not exploitative."¹³ Broadly speaking, arguments against Moral Impermissibility claim there are no other wrong-making properties that override Pareto and consent.

Yet, there seem to be *many* conditions beyond Pareto and consent that make transactions wrongful, and thus, many reasons to reject Pareto Permissibility. Philosophers have variously argued that transactions are impermissible when they are disrespectful, when they involve trade in non-commodifiable goods, when they violate moral rights, and (most applicable to the present discussion) when they are unfair. Clearly we need not look far for reasons to reject Pareto Permissibility. When considering exploitation, a concern for fairness stands out as an obvious wrong-making property. Indeed, most theories of exploitation begin with the claim that exploitation involves taking *unfair* advantage.¹⁴ Consider the following in place of Pareto Permissibility:

Necessary Fairness Transactions are permissible only if they are *fair*.

¹¹ Some sweatshops may lack this kind of consent; however, it is relatively easy to explain the wrong of this kind of work: if workers were not suitably informed about the conditions of their contract, then the transaction is best described as fraud.

¹² Kristof (2009, 14D). Note that Kristof's claim is not that sweatshops ought to pay *lower* wages, but rather that there ought to be more sweatshops.

¹³ Zwolinski (2007, 711). See also Powell and Zwolinski (2012). More generally, Pareto Permissibility may be attractive to those who endorse the welfarist thesis that the moral goodness of a state of affairs is an increasing function of the welfare in that state.

¹⁴ See Elster (1982), Goodin (1987), Sample (2003), Steiner (1984), and Wertheimer (1996).

Necessary Fairness claims that at least one reason transactions may be impermissible is that they are unfair. Note that the condition is much weaker than Pareto Permissibility because it makes no claim about what conditions are jointly sufficient for transactions to be permissible. For example, Necessary Fairness allows that consent, Pareto improvement, and fairness are jointly sufficient for permissibility. And it allows for the addition of other concerns to this list (and for the removal of both consent and Pareto). Though existing accounts of exploitation disagree about how the unfairness in exploitative transactions should be explicated, there is significant agreement that conditions beyond strict Pareto and consent—which, because they pertain to distributive concerns, fall broadly under the label ‘fairness’—must be satisfied in order to ensure the permissibility of transactions.¹⁵ Unfairness is one reason for the intuition captured by Moral Impermissibility that some Pareto improving and consensual transactions, such as sweatshop labour, are nevertheless impermissible. Consequently, unfairness also provides a reason to reject Pareto Permissibility. Indeed, this appeal to fairness is so obvious that we may wonder what motivates the defenders of Pareto Permissibility.

One charitable response to Pareto Permissibility is that those drawn to the condition really intend to defend a closely related claim, similar to what Alan Wertheimer has called ‘non-worseness’:¹⁶

Non-worseness If a transaction is strictly Pareto improving and consensual, then engaging in the transaction is morally *better* than not transacting.¹⁷

Although similar to Pareto Permissibility, Non-worseness does not conflict with Necessary Fairness. It allows that transactions can be strictly Pareto improving and consensual, yet impermissible because unfair. Non-worseness claims that, provided Pareto and consent are satisfied, transacting, though perhaps impermissible, is nevertheless better than not transacting. The endorsement of both Necessary

¹⁵ For the purpose of the paradox that follows, it is unnecessary to specify a particular conception of fairness, though the resolution of the paradox I outline in Sect. 4 places some broad constraints on conceptions of fairness.

¹⁶ Wertheimer offers his own analysis of non-worseness in Wertheimer (1996, 289–293).

¹⁷ Wertheimer states his similar claim as follows: “Given that I have a right not to transact with B and that transacting with B is not worse than not transacting with B, it can’t be seriously wrong for me to engage in an unfair transaction with B” (Wertheimer 1996, 289). Wertheimer’s Non-worseness is not equivalent to the condition I outline above. The negation of ‘worse than’ is ‘better than or equal to thus, Non-worseness in Wertheimer’s use implies that Pareto and consent are sufficient for transacting to be better than or equal to not transacting.

Also note that Non-worseness introduces the relation ‘morally better than’, or ‘ \succ^M ’. This relation is defined over the nonempty set $\mathcal{T} = \{t_0, t_1, \dots, t_T\}$ of possible transactions (where non-transaction is denoted by t_0), such that for any t_i and $t_k \in \mathcal{T}$, $t_i \geq^M t_k$ iff t_k is morally at least as good as t_i ; and $t_i \sim^M t_k$ iff $t_i \geq^M t_k$ and $t_k \geq^M t_i$; and finally $t_i \succ^M t_k$ iff $t_i \geq^M t_k$ and $\neg(t_i \sim^M t_k)$. I shall assume this moral betterness relation is reflexive, transitive, and complete and so forms a total moral goodness order. Thus, more formally, non-worseness claims that for any $t_T \in \mathcal{T}$, if t_T is strictly Pareto improving and consensual, then $t_T \succ^M t_0$.

Finally, note that the transitivity assumption in the moral goodness order—in (Temkin 2012)—and completeness assumption—in (Vallentyne 1993)—have been challenged, though as Voorhoeve (2013) points out the transitivity challenge is not wholly convincing. A discussion of the completeness assumption though interesting would, I fear, be beyond the scope of this paper.

Fairness and Non-worseness is an attractive resolution of the initial tension because this approach captures the idea that there is something good about Pareto improving, consensual transactions—they are better than not transacting—and yet, it is consistent with the intuition that some of these transactions, though better than not transacting, may nevertheless be impermissible.

Indeed, both conditions lie at the heart of our thinking about cases of exploitation. Exploitation involves a fundamental concern for the welfare of the exploited, but this concern differs from a general concern that their welfare is low. Rather, it is a concern that the welfare of the exploited is not being increased *enough* in a transaction. Though a particular transaction results in a welfare gain for the exploited, the claim that they are exploited entails that they *ought to* receive more. Thus, our concerns about exploitation involve both an intuition that the gains from transacting should be *fairly* distributed and an intuition that increases in welfare are morally valuable. Unfortunately, though Necessary Fairness and Non-worseness are individually compelling, jointly satisfiable, and essential to our thinking about exploitation, when combined with two compelling additional claims they generate a paradox.

2 The Paradox

Suppose Bob wants to purchase an apple from Alice and that the most he is willing to pay for the apple is \$2.50. Also suppose Alice's reservation price is \$1; at any price below \$1 Alice would rather keep the apple than sell it.¹⁸ Together these suppositions mean that any transaction that prices the apple above \$1 and below \$2.50 will be strictly Pareto improving. That is, at any price in that range both Alice and Bob will strictly prefer transacting to not transacting. Suppose also that the most plausible account of fair transaction (which, for the time I need not specify) entails that the fair price for the apple is \$1.50. In order to focus on the issue of fairness I will assume—both here and throughout—that the transaction does not violate any other possible restrictions on the permissibility of transactions. I will limit my attention to cases where fairness is the only 'permissibility variable' in play. Finally, suppose that Alice offers to sell Bob an apple for the *unfair* price of \$2.

Assuming Bob wants to minimise the price he pays for an apple, Bob will prefer (\succ_B)¹⁹ the fair price of \$1.50 to the unfair price of \$2, since \$1.50 is less than \$2. Bob will also prefer the unfair price to not transacting, since the unfair price of \$2 remains below his reservation price of \$2.50 is:

¹⁸ Alice's reservation price may, but need not, be a function of the cost of producing the apple. If it is, then when Alice sells the apple for less than \$1 she makes a loss. But she may have other reasons for setting her reservation price at \$1.

¹⁹ The preference relation \succ is defined as follows. Let $\mathcal{N} = \{1, 2, \dots, N\}$ be the set of agents (where $N > 1$) and assume these agents have preferences over the transactions in \mathcal{T} , such that, for any $j \in \mathcal{N}$, the binary relation \succsim_j is a preference relation defined over \mathcal{T} , such that, for any t_i and $t_k \in \mathcal{T}$, $t_i \succsim_j t_k$ iff agent j prefers t_i at least as much as t_k ; and $t_i \sim_j t_k$ iff $t_i \succsim_j t_k$ and $t_k \succsim_j t_i$; and finally $t_i \succ_j t_k$ iff $t_i \succsim_j t_k$ and $\neg(t_i \sim_j t_k)$. I make the standard assumption that the agents' preferences over transactions are reflexive, transitive, and complete and can therefore form a total order.

$$\text{Fair } (\$1.50) \succ_B \text{Unfair } (\$2) \succ_B \text{No transaction} \quad (1)$$

Now consider Alice. Assuming Alice wants to maximise her revenue, she will prefer (\succ_A) the unfair price (\$2) to the fair price (\$1.50) since \$2 is greater than \$1.50. She also prefers the fair price to not transacting, since \$1.50 is above her reservation price. Thus, Alice's preference order is:

$$\text{Unfair } (\$2) \succ_A \text{Fair } (\$1.50) \succ_A \text{No transaction} \quad (2)$$

Although Alice and Bob have opposing preferences between fair and unfair prices, both prefer some form of transaction to no transaction. If both also consent to the transaction, then it seems that transacting at any price in the range between \$1 and \$2.50 is not only prudentially better than not transacting, but it is also morally better. That is, as Non-worseness claims, it is morally better to transact unfairly (at \$2) than to not transact, since strict Pareto and consent are satisfied. With a particular case before us, it is difficult to argue against Non-worseness. After all, both Alice and Bob would prefer transacting at \$2 to not transacting, both consent to this transaction, and—by Pareto—there are no negative externalities involved in the transaction. In fact, it seems that the only harm would come from *preventing* the transaction. Those who would deny Non-worseness owe an account of why proscribing transaction in the name of some other value should override the transactors' preferences and welfare. Absent an explanation, the denial of Non-worseness appears perniciously moralistic.

As I noted, Non-worseness captures concerns similar to those expressed by Pareto Permissibility, but unlike Pareto Permissibility it is consistent with Necessary Fairness. However, one difference between Non-worseness and Pareto Permissibility is that the betterness relation expressed in Non-worseness—the claim that transacting is 'morally better than' not transacting when Pareto and consent are satisfied—is an *evaluative* binary relation that provides an ordinal ranking of options, while the relation expressed in Necessary Fairness (and Pareto Permissibility) is a *deontically modal* concept that partitions acts into the two jointly exhaustive and mutually exclusive categories of 'permissible' and 'impermissible'.²⁰ While one act can be better than another, it cannot be more permissible than another permissible act. Unlike goodness, permissibility does not come in degrees.²¹

Relating evaluative and deontic judgements is difficult. An act's permissibility does not, generally, give us any information about the act's goodness, and vice versa. However, there is one plausible exception to this claim. It seems that permissible acts are better than impermissible acts.²² That is,

²⁰ Here, and throughout, I will assume permissible and impermissible are mutually exclusive and exhaustive.

²¹ It might be argued that one can make sense of one act being 'more permissible' than another. Insofar as this claim makes sense, it seems that it is because it references the goodness of the 'more permissible' act. Peterson (2013) has defended the notion of 'degrees of rightness', though his is a minority view.

²² Some non-consequentialists may reject this claim. However, as I discuss in section 3.1 there are good reasons to ensure our normative theories respect it.

Betterness of Permissibility For any two acts x and y , if x is morally permissible and y is impermissible, then x is morally better than y .²³

For example, suppose that theft is impermissible, and murder is morally worse than theft. Betterness of Permissibility implies that murder is also impermissible. The condition stipulates that impermissible acts do not suddenly ‘flip’ and become permissible as the goodness of the acts in question diminish.

Recall the earlier example of the apple sale. By hypothesis, the price of \$1.50 for the apple is fair and consensual, though all prices above \$1 and below \$2.50 are strictly Pareto improving. Since the transfer of the apple for \$1.50 is fair, consensual, strictly Pareto improving (and was assumed to satisfy other potential permissibility conditions), it is permissible. However, because the \$2 price does not satisfy the fairness condition it is impermissible, according to the Necessary Fairness premise. If permissible acts are morally better than impermissible acts, then selling the apple for \$1.50 is better than selling the apple for \$2. That is,

$$\text{Fair } (\$1.50) >^M \text{Unfair } (\$2) \quad (3)$$

In addition, Non-worseness claims that if a transaction is strictly Pareto improving and consensual, then engaging in the transaction is better than not transacting. Since transacting at \$2 satisfies these conditions, transacting at \$2 is morally better than not transacting. So, by Non-worseness,

$$\text{Unfair } (\$2) >^M \text{No transaction} \quad (4)$$

The combination of these two orders gives us the following goodness order for the three options:

²³ I believe Betterness of Permissibility is compelling as a broad metaethical constraint and thus, it takes as its scope the set of all acts; however, one need not accept such a broad scope to generate the paradox explored in this paper. The paradox is generated when the scope of the condition is limited to the transactions in \mathcal{T} . Formally, Betterness of Permissibility states that for any $t \in \mathcal{T}$ if t_m is morally permissible and t_n is morally impermissible then $t_m \succ^M t_n$.

Betterness of Permissibility (BP) is a stronger version of a related condition, Monotonicity of Impermissibility (MI), which states that acts worse than impermissible acts are also impermissible. BP implies MI, but MI does not imply BP.

Proof:

BP \rightarrow MI: if MI is violated, then so is BP. Suppose, violating MI, an act worse than an impermissible act is permissible: it is impermissible to x , permissible to y and y is morally worse than x . Then there is a permissible act, y , that is not better than an impermissible act, x .

$\neg(\text{MI} \rightarrow \text{BP})$: if BP is violated, then MI can still be satisfied. BP is false in degenerative cases where all acts have equal value, such as when x is permissible, y is impermissible and x is equally as good as y . But here MI is not violated, since MI says nothing about the impermissibility of acts that are equally as good as an impermissible act. \square

The Paradox can be generated by appealing to the weaker MI; however, the ease of interpretation of BP make the stronger condition much more straightforward. Also note that, though the order in (3) is inessential for the paradox, it depends on the stronger BP condition.

$$\text{Fair } (\$1.50) >^M \text{Unfair } (\$2) >^M \text{No transaction} \quad (5)$$

Order (5) can explain the protracted disagreement between those who would condemn sweatshops on grounds of exploitation and those, like Kristof, who argue that we need more sweatshops. In a sense, both are correct. Kristof, who argues it is better for Christina to be employed by Chentex rather than unemployed can endorse Non-worseness, which gives order (4). Those who condemn sweatshops can endorse the Necessary Fairness condition, which—along with the Betterness of Permissibility—gives order (3). Both parties can endorse order (5), which claims that though employment is morally better than unemployment, employment at a fair wage is better still.

We now have two attractive results. First, as I noted above, Non-worseness is compatible with Necessary Fairness. This compatibility allows us to claim both that there is something good about Pareto improving, consensual transactions and that these transactions may nevertheless be impermissible. Second, Non-worseness, Necessary Fairness, and Betterness of Permissibility can generate order (5), which resolves the disagreement between sweatshop defenders and critics: though it is true that sweatshops are better than nothing, they are not better than situations where the terms of the labour contract are fair. Thus, the three conditions not only capture the concerns about fairness and welfare present in our thinking about exploitation, they also resolve the debate that motivates the initial tension outlined in Sect. 1.

Although these two results are attractive, a problem looms. Taken together, Non-worseness, Necessary Fairness, and Betterness of Permissibility conflict with the following claim:

The Optionality of Ordinary Transactions Strictly Pareto improving and consensual transactions are not *ordinarily* obligatory.

The ‘ordinarily’ clause in the condition restricts its scope to day-to-day transactions that do not feature extra-ordinary moral obligations. For example, Alice may have an obligation to purchase Bob’s car if she promised to do so. Promises, reparations, and especially duress can make certain transactions obligatory. The Optionality premise does not deny these facts. Rather, it claims that absent other morally obliging circumstances, we are not obliged to transact simply because a transaction would be mutually beneficial and consensual.²⁴ Or, what amounts to the same thing, the condition claims that ‘ordinary’ cases it is *permissible* to not transact despite the fact that doing so would be mutually beneficial. The optionality condition denies that transactions become obligatory simply in virtue of being strictly Pareto improving.²⁵

²⁴ Note that ‘ordinary’ is not used in the sense of ‘commonplace’ but rather in a special sense of ‘morally ordinary’. So, extreme poverty may indeed be ordinary in the sense of being commonplace, but extreme poverty also provides reasons that we may plausibly be obliged to transact with others—for example through charitable donations—and so it is not a morally ordinary situation.

²⁵ It is important to note that if Bob fails to engage in a Pareto improving and consensual transaction with Alice he is *irrational*. While the claim that the set of moral acts must be a subset of rational acts is compelling, the claim that the two sets are identical seems implausibly strong. Indeed, it is likely that many forms of consumer behaviour are irrational, but not impermissible. See Tversky and Kahneman (1981).

The conflict between the Optionality premise and the conjunction of Non-worseness, Necessary Fairness, and the Betterness of Permissibility occurs in the following way. According to Necessary Fairness, fairness is necessary for the permissibility of a transaction. In the apple sale case, because the \$2 transaction is (by hypothesis) unfair it is impermissible. Non-worseness claims that if a transaction is strictly Pareto improving and consensual, it is morally better than not transacting. Thus, as expressed in order (4), not transacting is morally worse than engaging in the \$2 transaction. Betterness of Permissibility implies that if an act is worse than an impermissible act, then it too is impermissible. Since (by Non-worseness) not transacting is worse than transacting unfairly, and (by Necessary Fairness) transacting unfairly is impermissible, then (by Betterness of Permissibility) not transacting is also impermissible. Yet, this result conflicts with the Optionality premise.²⁶ Surely Bob does nothing wrong when he fails to purchase an apple from Alice. Either the Optionality premise, or one of the three premises entailing the *obligation* of ordinary transactions must be abandoned.²⁷

3 Evaluating the Conditions

Non-worseness captures the common consequentialist concern to promote individual welfare. Necessary Fairness captures the non-consequentialist idea that those who transact unfairly do something impermissible. These conditions—which are compatible—capture concerns for welfare and fairness that shape our thinking about exploitative transactions. Not only do they capture the moral concerns that feature in exploitation, as we saw in the previous section, they also help explain the disagreement between defenders and critics of sweatshops.

The remaining two conditions—Betterness of Permissibility and Optionality of Ordinary Transactions—can be interpreted as conditions that respectively prohibit undesirable features of non-consequentialist and consequentialist moral frameworks. Optionality says that engaging in Pareto improving and consensual transactions should not be morally required. If we endorse both Optionality and

²⁶ Here I am assuming that $(\Diamond(x \wedge \neg x) = \neg \Box x)$. That is, the Optionality condition's claim that it is optional to transact implies that it is not obligatory to transact, which contradicts the implication of the first three conditions that it is obligatory to transact.

²⁷ It can be shown that the paradox also arises when Wertheimer's form of Non-worseness (Pareto and consent are sufficient for transacting to be better than *or equal to* not transacting) replaces the version defined above and when Monotonicity of Impermissibility (MI) replaces Betterness of Permissibility. However if *both* alternative conditions are used, then these conditions along with Necessary Fairness do not contradict the Optionality condition.

If transacting unfairly at \$2 is impermissible, then this is, by Wertheimer's Non-worseness condition, morally better than or equal to not transacting. If it is equivalent, then MI does not imply that not transacting is impermissible. MI only states that an act that is *worse* than an impermissible act is also impermissible. Thus, when both weaker conditions are adopted there is no paradox. But this is an unattractive solution. The paradox can be reinstated with the addition of the rather weak claim that there can be no difference in permissibility without a difference in moral value (I discuss the problematic consequences of denying this claim in the following section). With this added condition, if transacting unfairly at \$2 is impermissible and not transacting is morally equivalent in value to transacting at \$2, then not transacting is also impermissible, contradicting the Optionality condition.

Non-worseness, then this is true even if doing so is morally better. In this sense, Optionality can be seen as a condition that allows for supererogation. Or put in another way, Optionality rules out a form of the demandingness challenge to consequentialism, which claims the obligation to perform the act that brings about the most good is too demanding.

Each of the conditions is initially appealing, yet they cannot all be true. Consistency is regained, and the paradox resolved by simply abandoning one of the four premises. Yet, while dropping one of the premises removes the logical conflict between the four claims, dropping a premise is not, on its own, a *solution* to the paradox. A full solution is a rejection–justification pair that not only identifies a condition to abandon, but also justifies this rejection by explaining why, though initially attractive, the condition is nevertheless false.

First consider a rejection of the Optionality condition on utilitarian grounds. Utilitarians might endorse order (5) by arguing that because the only morally permissible act is that which brings about the most moral goodness and, since by Non-worseness not transacting is worse than unfair (and fair) transactions, not transacting is impermissible. Of course, this approach faces considerable challenges. For example, by making all Pareto improving transactions morally obligatory, the response is very morally demanding.²⁸ More problematic though, is that this approach faces a ‘moral phenomenology’ problem, by failing to account for our moral attitudes about cases of exploitation, such as sweatshops. By accepting that not transacting is morally worse than transacting unfairly this response would imply that those who are uninvolved with sweatshops perform morally worse acts than sweatshop owners. Furthermore, it seems that those who adopt this approach must grant that consequentialism simply cannot account for the wrongs of exploitation, which, of course, in itself would be an interesting result.

An alternative solution that falls within the consequentialist framework would be to justify a rejection of Necessary Fairness by appealing to welfarism, the view that the goodness of states is based only on individual welfare (or, sometimes more narrowly, preference satisfaction). Welfarists cannot endorse concerns for fairness that are independent from individuals’ preferences for fairness. Instead, any concern for fairness must be ‘written into’ the individual transactors’ preferences: insofar as Alice or Bob have a concern for fair transactions their reservation prices and preference profiles will take this concern into account. Yet, like the previous solution, this approach would rank exploitative transactions above non-transaction and so has a problem accounting for certain aspects of moral practice. Perhaps a stronger reason for rejecting this solution is that an important group of individuals lack an overriding concern for fairness: the very poor. When persons’ welfare is very low they are likely to attach a much greater weight to whatever material gains they can acquire. As our level of material possessions increases, the value we attach to each additional material possession decreases. Consequently, if we have no, or few, material possessions, material gain is likely to outweigh a gain in non-material

²⁸ Though, consequentialists have developed responses that diminish the force of demandingness objections—see Sobel (2007) for such an account—these responses do not seem to apply to the kinds of demands entailed by the rejection of the optionality condition.

goods, such as fairness. And this means that the very poor will rarely, if ever, be able to give much weight to fairness when they transact. But if we follow welfarists in claiming that moral goodness increases only with welfare (or preference satisfaction), then we cannot condemn unfair transactions with the very poor. This is appalling—when someone is in such a bad state that they lack the ability to protest by refusing to transact, we should not conclude from their overriding preference for material gain that any Pareto improving and consensual transaction is permissible. Still, like the previous solution, the welfarist solution is also independently interesting since it would entail not only that there is no independent sense of fairness that can be used to condemn transactions as exploitative.

Still, it is perhaps not surprising that welfarism, utilitarianism, and similar forms of consequentialism are not *straightforwardly* able to account for the impermissibility of exploitation. Most persons concerned about exploitation assume that it is impermissible (accepting Necessary Fairness) and that the obligation to not exploit is a conditional obligation that we inherit through our interaction with those with whom we transact (accepting Optionality). They would endorse the claim that though we are under no obligation to transact with others, if we choose to do so, we ought to do so fairly. However, it is difficult to square this claim with the concern for welfare enshrined in Non-worseness. The only remaining options for those who accept the conditional obligation are to reject Betterness of Permissibility or Non-worseness.

The appeal of Betterness of Permissibility may not be initially clear. The condition specifies a relation between evaluative claims about the goodness of a state of affairs and deontic claims about the rightness of acts. Such ‘bridge principles’ are controversial and, indeed, the Betterness of Permissibility condition can be violated within some non-consequentialist moral frameworks. However, Betterness of Permissibility can be interpreted as a principle that prohibits the paradox of deontology, and thus, may be interpreted as a consistency requirement for a moral framework.

The connection here can be illustrated by considering the two cases in which Betterness of Permissibility is violated. First, note that in order for Betterness of Permissibility to be directly violated within a moral framework—rather than merely inexpressible—the framework must make claims about both rightness and goodness.²⁹ The ability to explain not only what acts are permissible or impermissible, but also an explanation of why some acts are worse (or better) than others is a desideratum of any complete moral framework. We want moral theories that allow both for moral evaluation and that are action guiding.

A moral framework violates Betterness of Permissibility whenever an act lacks the property (or properties) that, according to the theory, makes an act permissible, but has more of the property (or properties) that, according to the theory, make acts morally good. In other words, Betterness of Permissibility is false when there is an impermissible act that is better than a permissible act. Though violations of the condition are possible for frameworks that include both accounts of goodness and

²⁹ For example, some scalar consequentialists, such as Norcross (2006) argue we should abandon the notion of rightness.

permissibility, such violations encounter a dilemma: If a moral framework violates Betterness of Permissibility, then it is either ‘axiologically incongruous’ or it encounters the ‘paradox of deontology’.

Consider the first horn. If for a particular framework the properties that determine the goodness of a state are wholly independent from the properties that determine the rightness of an act, then within that framework doing the right thing will have nothing to do with what—*according to that same framework*—makes the world a better place. If the properties that determine rightness and goodness are unrelated then the moral framework suffers from a kind of ‘two-mindedness’ about moral value, it is, in this sense, axiologically incongruous. Further, if goodness and rightness are not connected, then right acts cannot be justified by appealing to their *value*. Yet, without such a connection it seems very hard to explain or justify an account of rightness to others who do not accept it. This fork of the dilemma should be resisted.

In order to avoid this problem, the properties that determine rightness should be tied in some way to the properties that determine goodness. That is, a reason impermissible acts are prohibited is that they make the world a worse place. This leads us to the second horn. If the properties that determine rightness and goodness are tied together in this way, then it seems inconsistent to claim that a concern for rightness justifies prohibiting act ψ if the performance of act ψ will bring about more of the property that the concern for rightness is intended to protect or promote. In other words, why not violate a restriction whose point it is to promote a property if doing so leads to *more* of that property in the world? Principles that prohibit what they are designed to promote suffer from a form of inconsistency. This problem is one form of the well known paradox of deontology.³⁰

Note that there is a second way Betterness of Permissibility can be violated. The condition is false when a permissible act and an impermissible act are equally good. Here again we face a dilemma. Suppose goodness and rightness are related. Then in this case, the two cannot be equivalent, for if they differ in rightness and rightness contributes to goodness, then they cannot be equally good. The situation is definitionally impossible. Suppose that rightness does not contribute to goodness. Then, though there is no definitional inconsistency, we again face axiological incongruity: why we should care about the right making property since there is, by supposition, no moral value for the property.

In response to the dilemmas that arise from the violations of Betterness of Permissibility, non-consequentialists may claim that what makes an act good is part—but not all—of what makes the act right. That is, the right-making properties are part of, but not identical to, the good-making properties. In general, there may be many different properties that are relevant for rightness and many different properties relevant for goodness. However, these mixed accounts do not avoid the dilemma. For any property apart from goodness that could be relevant to rightness we face the question, ‘why does this property matter for rightness?’ Any answer that says this property is right-making, but not associated with goodness is axiologically incongruous. And any answer that ties the property to goodness faces the paradox of

³⁰ See, for example, Nozick (1974), Scheffler (1988), and Heuer (2011).

deontology if it violates the Betterness condition in the first way and cannot arise if it violates the condition in the second way. Thus, the dilemma plagues both ‘pure’ and ‘mixed’ accounts.

If a moral framework violates the Betterness of Permissibility by claiming impermissible acts are better than permissible acts, then it faces either the paradox of deontology or it is axiologically incongruous. If a moral framework violates the condition by claiming a permissible act and an impermissible act are equally good, then the two conditions must be unrelated and, again the framework is axiologically incongruous. Finally, any ‘mixed’ account of the relationship between goodness and rightness can be decomposed into a ‘pure’ account where these problems re-emerge. Assuming non-consequentialists would rather confront the paradox of deontology than axiological incongruity, we may interpret the Betterness of Permissibility as a consistency requirement that avoids the paradox of deontology.

In summary, the first two conditions, Non-worseness and Necessary Fairness respectively capture consequentialist and non-consequentialist concerns present in our thinking about exploitation and they can explain ongoing disagreement about cases of exploitation. The Optionality and Betterness conditions can be seen as intuitions that our analysis of exploitation should be able to account for supererogation and avoid the paradox of deontology. Thus, the four conditions that comprise the paradox ask us to capture both consequentialist and non-consequentialist concerns while avoiding the well-known problems that confront each moral framework. Seen in this light, the fact that these four conditions are jointly inconsistent is perhaps unsurprising. In the next section I outline a solution to the paradox that attempts to balance a concern for welfare against a concern for fairness and in so doing, retains the spirit of all four conditions.

4 Resolving the Paradox

If one accepts the arguments in favour of retaining Betterness of Permissibility and Necessary Fairness and wishes to avoid biting the bullet by denying the Optionality condition, then only one option remains: Non-worseness. In the sections that follow I will argue that the paradox of exploitation should be resolved through a weakening of the Non-worseness condition, combined with a clarification of what, precisely, it means for a transaction to be unfair.

Recall the original Non-worseness condition: ‘if a transaction is strictly Pareto improving and consensual, then transacting is better than not transacting’. I argue that Non-worseness is compelling only when weakened:

Weak Non-worseness If one or more strictly Pareto improving and consensual transactions are possible, then for at least one of these transactions, transacting is morally better than not transacting.

Consider the following goodness order:

$$\text{Fair } (\$1.50) >^M \text{No transaction} >^M \text{Unfair } (\$2) \quad (6)$$

This order violates the original Non-worseness condition because the unfair transaction, though Pareto improving and consensual, is ranked below non-transaction. However, (6) respects Weak Non-worseness since, of the two possible Pareto improving and consensual transactions, at least one is morally better than not transacting.

Not only does (6) satisfy Weak Non-worseness, it is also consistent with the remaining three conditions. By Necessary Fairness, unfair transactions are impermissible and by the Optionality of Ordinary transactions, not transacting is permissible. In (6), not transacting is morally better than unfair transaction. Because both permissible acts are ranked above the impermissible act of unfair transaction, order (6) also respects the Betterness of Permissibility. Therefore, when the original Non-worseness condition is replaced by Weak Non-worseness, the new condition is jointly consistent with Necessary Fairness, Betterness of Permissibility, and the Optionality of Ordinary Transactions.

Though Weak Non-worseness is consistent with the remaining three conditions in order (6) there is a special class of cases in which it *appears* a conflict might still arise: when *any* possible Pareto improving transaction is unfair. Suppose there are two (and only two) possible Pareto improving transactions in the apple sale case, both of which are unfair. For example, suppose the fair price for the apple is again \$1.50, but the only possible strictly Pareto improving transactions between Alice and Bob are an apple for (the unfair price of) \$2 or an apple for (the unfair price of) \$1.75. According to Weak Non-worseness *at least one* of these transactions must be better than not transacting. Perhaps it is \$1.75, the price that is closest to the fair \$1.50 transaction. In this case, by Weak Non-worseness, transacting at \$1.75 is better than not transacting:

$$\text{Unfair } (\$1.75) >^M \text{No transaction} \quad (7)$$

But (7) violates Betterness of Permissibility. By Necessary Fairness transacting at \$1.75 is impermissible because it is unfair, which means that an impermissible act is better than a permissible act. Does this mean that Weak Non-worseness fails to solve the paradox in all cases?

No, it does not. At least some point in the interval between Alice's reservation price and Bob's reservation price must be fair. That is, if any strictly Pareto improving transaction is possible, then at least one such transaction must also be fair. If one of \$1.75 or \$2 is fair, then, in the above example Weak Non-worseness does not lead to a conflict with Betterness of Permissibility. In general, this means that the special class of cases in which Weak Non-worseness fails cannot arise.

But why should we believe that if there is any possible, and permissible, strictly Pareto improving transaction, then at least one such transaction must be fair? In order to see that it is, we must distinguish fairness as a distributive concept from fairness as procedural concept. The kind of unfairness we are concerned with when we say that exploitation involves taking unfair advantage is distributive unfairness;

that is, in exploitations the distribution of gains from transacting is a *maldistribution*. These gains ought to have been distributed in some other way. What a person gains from transacting is the difference between what they get from engaging in the transaction from what they get if they refrained from transacting. If Bob would have paid up to \$2.50 for the apple, then when he pays anything less than \$2.50, he gains by transacting. Similarly, if Alice would have sold the apple for any price greater than \$1, then when she receives anything more than \$1, she gains from transacting. When Alice and Bob transact with each other, they both gain at any price above \$1 and below \$2.50. What we are concerned with in cases of exploitation is how this mutual gain should be fairly divided—with how the cake should be cut. In this case, the fair division must be some division of the cake itself, that is, of the gains from transacting. Thus, a fair distribution in the context of transacting must lie somewhere in the interval between each individual's reservation price.

Transactions may, of course, be unfair for other reasons. For example, suppose a transaction in which Alice and Bob exchange an apple for \$1.50 *distributively fair*, but suppose also that Bob cut in front of Carol in the queue and bought the last apple, which Carol intended to buy. Here we may also describe the transaction as 'unfair', not because the distribution of the gains between the buyer and seller was a maldistribution, but because one buyer was unfairly excluded from the market. The transaction is *procedurally* unfair.

These two forms of unfairness are often linked—in many cases procedural unfairness may cause distributive unfairness. Suppose Carol would have paid \$2 for Alice's apple, but because she was prohibited from entering the market, Alice only received \$1 for her apple from Bob. In this case, both forms of unfairness are present. Carol's being barred from the market is procedurally unfair and that leads to an unfair distribution of the gains from transaction between Alice and Bob. However, when we claim that Nicaraguans are exploited by Chentex this is not a claim about *procedural* unfairness. It is a claim about distributive unfairness, a claim that they ought to receive more for their labour.

Yet, even if we restrict our attention to distributive fairness, it seems that there can be cases where the fair transaction is not found in the interval between the transactors' reservation prices. Suppose Alice steals Bob's car and offers to sell it back to him. If Bob has no choice but to pay Alice to get the car back (and presuming he wants to replace it), then he will be willing to pay up to the amount that replacing the car would cost—if a new car costs \$500 (and if Bob attaches no sentimental value to the car), Bob would be willing to pay Alice up to \$499. Suppose that Alice doesn't need to use the car herself, and so doesn't attach much use-value to the vehicle, but she does know she can sell the car elsewhere for \$100 and so will sell the car to Bob for anything above this amount. If it is true that the distributively fair price must lie between the transactors' reservation prices, then the fair price for the car must be between \$101 and \$499. Surely this is incorrect. The fair 'price' for the car is \$0.

The act that leads to the morally best state of affairs is one in which Alice simply returns Bob's car. Since Bob values the car more than any price below \$500, then it seems that the next best options are those in which Bob gives amounts less than

\$500 to Alice in exchange for his car (e.g., \$1 for the car is better than \$2, is better than \$3, etc.). And finally the morally worst ranked option is one in which there is no transaction (here Bob's loss is equivalent to a loss of \$500).

$$\text{Return car} >^M \text{Transactions } \$1 - \$499 >^M \text{No transaction} \quad (8)$$

Order (8) captures our intuitions about the goodness of alternatives in the car case. Here the fair transaction is the car's free return. Because transactions \$1–\$499 are unfair, they are impermissible by Necessary Fairness. Weak Non-worseness is satisfied since at least one Pareto improving and consensual transaction (transactions \$1–\$499) is morally better than not transacting. But as before, Betterness of Permissibility implies that not transacting is impermissible, contradicting the Optionality of Ordinary transactions. However, in this case the optionality condition is not compelling. Transacting is not morally optional because Alice is *obliged* to engage in the only fair transaction—she is obliged to return the car.

There is an important difference between the car case and the apple case. In the car case, a prior injustice—the car's theft—makes transacting obligatory and, therefore, the reason the Optionality of Ordinary transactions is implausible in this context is that the 'ordinary' clause is not satisfied. The transaction is not an 'ordinary' transaction because Alice is obliged to pay reparations for the theft. The car case lies outside the domain of transactions in which the paradox arises. This means that my prior claim that the fair transaction must lie in the interval between the transactors' reservation prices is true only if there are no extra-ordinary obligations to transact.

'But,' it might be objected, 'aren't some exploitations like this—in some cases aren't those who are poor being sold something they are entitled to receive for free, as reparation for past injustice?' The answer is yes, some transactions ordinarily described as exploitations fall into this category. And in these cases, the fair and obligatory transaction will be one in which the exploited simply receive what they are owed without exchanging anything in return. The fair transaction will be *unilateral* and thus, not a Pareto improvement. But the question that the paradox poses is whether *absent* any extra-ordinary obligation, transactions become obligatory in virtue of being Pareto improving. The original Non-worseness, Necessary Fairness, and Betterness of Permissibility conditions imply that they do, contradicting the ordinary intuition, captured by Optionality of Ordinary Transactions, that they do not.

'Surely,' the objection continues, 'if we limit our conception of fairness to distributive fairness *and* we limit our domain to those in which there is no obligation to engage in a non-Pareto improving transaction, then the claim that the fair transaction must be some *distribution* of the gains from transacting—that it must be found in the interval between the transactors' reservation prices—is true by definition.' Yes. But this is precisely the point. Because exploitation is a matter of distributive unfairness (which may, nevertheless, be caused by procedural unfairness) and because we are attempting to determine whether Pareto itself can entail an obligation to transact, we must limit the conception and domain of fairness in this way. And given these restrictions, it follows, as I claimed above, that if there

is any possible, and permissible, strictly Pareto improving transaction, then at least one such transaction must be fair.

So, to return to the issue that motivated this discussion of fairness, if the fair transaction must be a strictly Pareto improving transaction—if it must fall between the transactors' reservation prices—then this means that it is not possible for cases like that depicted in order (8) to arise. If such orders cannot arise, then there can be no case in which Weak Non-worseness fails to solve the paradox. Therefore, Weak Non-worseness, Necessary Fairness, Betterness of Permissibility *and* the Optionality of Ordinary Transactions can be jointly satisfied.

In Sect. 1 I claimed the fundamental concern in cases of exploitation is the welfare of the exploited is not being increased *enough* in a transaction. This concern for welfare is not allayed by morally obliging situations where no strictly Pareto improving transaction takes place, for here the would-be exploited have the lowest welfare. As Joan Robinson famously claimed, “the misery of being exploited ...is nothing compared to the misery of not being exploited at all.”³¹ The priority of a concern for welfare means that any criteria of fairness must morally permit a possibility for the would be exploited to permissibly engage in a transaction in which they better themselves. In other words, absent extra-ordinary moral obligations, the fair transaction must lie in the interval between the transactors' reservation prices.

Weak Non-worseness ensures that possible criteria of fairness can satisfy this constraint. Since it stipulates that at least one Pareto improving transaction must be morally better than not transacting, it also ensures that fair transaction can always be morally better than not transacting. Of course, the original Non-worseness condition can also satisfy this constraint since it makes the even stronger claim that *all* Pareto improving transactions must be morally better than not transacting. However, although welfare is the fundamental concern in exploitation it is not the only one. Our responses to cases of exploitation also imply a strong concern for unfairness. We want to capture the conditional obligation that exploitation presupposes: when more favourable terms are available and are morally obliged, but exploiters fail to transact on these terms, they engage in activity that is morally worse than simply failing to transact. Though the original Non-worseness condition does not allow for such a claim, Weak Non-worseness *does* because it allows some Pareto improving transactions to be morally worse than not transacting, a difference reflected in the shift from order (5) to order (6).

The conflict of intuitions in cases of exploitation and the appeal of Weak Non-worseness can be illustrated by considering these conflicting moral orders:

$$Unfair >^M No\ transaction \quad (9)$$

$$No\ transaction >^M Unfair \quad (10)$$

The first, given by Non-worseness,³² captures a concern for welfare. It claims it is morally better to gain from a transaction even if the benefits of the transaction are

³¹ Robinson (1962, 45).

³² Assuming the unfair transaction is also strictly Pareto and consensual, Non-worseness claims that this transaction is better than not transacting.

unfairly distributed. Though very appealing, the ordering fails to capture our moral attitudes towards exploiters, who we think behave worse than those who do not exploit. The second order, which is jointly implied by Necessary Fairness, Betterness of Permissibility, and the Optionality conditions³³ can account for these attitudes; however, the second order fails to respect the intuition that it is better to benefit someone than to not. Weak Non-worseness allows the second order, and so can capture moral attitudes about exploiters, but it allows this order only if at least one *other* beneficial transaction (that is better for the exploited) is also possible. In this way Weak Non-worseness combines the attractive features of both orders. The condition can capture a subtle contextual change that the original Non-worseness condition cannot: it is the presence of the fair, but unchosen option that allows for the claim that exploiters' choice to transact unfairly is morally worse than not transacting. In other words, for unfair transaction to be worse than not transacting it must be the case that exploiters could have done otherwise, but did not.

The intuition that unfair transactions are worse than non-transaction only when a fairer transaction is possible can be captured only by a condition capable of expressing a three place relation between the options. Because Weak Non-worseness expresses a three place relation it can capture the intuition the wrongfulness of exploitation depends crucially upon whether transactors could have done otherwise. Thus, Weak Non-worseness manages to incorporate a concern for fairness, but makes this concern subordinate to a concern for the welfare of the would-be exploited.

5 Conclusion

I began with two conflicting claims. The first—captured by Pareto Permissibility—is the idea that strict Pareto improvement and consent were jointly sufficient for a transaction to be permissible. The second—captured by Moral Impermissibility—denies that strict Pareto and consent are jointly sufficient for permissible transaction. Necessary Fairness provided a justification for the rejection of Pareto Permissibility; that is, one reason to believe strict Pareto and consent are not sufficient for permissible transaction is that transactions may also be impermissible when they are unfair. I suggested that the primary motivation for the rejected Pareto Permissibility condition was a concern for transactors' welfare, which can be captured by the Non-worseness condition.

Together, Non-worseness and Necessary Fairness capture a concern for fairness and a concern for welfare, both of which inform our ordinary intuitions about exploitation. Additionally, along with Betterness of Permissibility, the conditions allow for goodness order (5), which explains the debate between those who support sweatshops on welfarist grounds and those who criticise sweatshops for being exploitative.

³³ Optionality says N is permissible. Necessary Fairness says U is impermissible. By Betterness of Permissibility it follows that: No transaction $>^M$ Unfair.

Unfortunately, these three conditions also *conflict* with the common intuition that Pareto improvement does not entail a moral obligation to transact. This intuition is captured by the Optionality of Ordinary Transactions condition. In the previous section I argued that the best way to solve the paradox, while incorporating both a concern for fairness and a concern for welfare, is to replace the original Non-worseness condition with Weak Non-worseness.

Some people may find the weak Non-worseness solution unappealing, not because it fails to solve the paradox, or offers an implausible balance between fairness and welfare, but because their antecedent commitments conflict with one of the other conditions that comprise the paradox. The ability to identify those commitments responsible for opposing intuitions about exploitation is, I believe, the primary advantage of framing the topic as a paradox and indeed, the main contribution of this paper. By clarifying our commitments the paradox allows for a more focused discussion about the ethics of transactions. While I have argued Weak Non-worseness provides the best way to satisfy all of our conflicting intuitions, an interesting feature of the paradox is that the rejection of *any* of the premises tells us something novel about exploitation.

Acknowledgments Versions of this paper were presented at the MANCEPT workshop 'exploitation', the Slovak Academy of Sciences, ECAP 8 at the University of Bucharest, Utrecht University, and the University of Bayreuth. I thank those present for their helpful comments. I also thank Vuko Andric, Richard Bradley, Susanne Burri, Sebastian Koehler, Christian List, Alice Obrecht, Olivier Roy, Orri Stefánsson, Hillel Steiner, Martin van Hees, Alex Voorhoeve, and my colleagues at the University of Bayreuth for their valuable suggestions.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Cohen, G. (1995). *Self-ownership, freedom, and equality*. Cambridge: Cambridge University Press.
- Elster, J. (1982). Roemer versus Roemer: A comment on "New directions in the Marxian theory of exploitation and class". *Politics and Society*, 11(3), 363–373.
- Goodin, R. (1987). Exploiting a situation and exploiting a person. In A. Reeve (Ed.), *Modern theories of exploitation* (pp. 166–197). London: Sage.
- Heuer, U. (2011). The paradox of deontology, revisited. In M. Timmons (Ed.), *Oxford studies in normative ethics* (pp. 236–267). Oxford: Oxford University Press.
- Kristof, N. (2009). Where sweatshops are a dream. *New York Times*, 2009, 14D.
- Norcross, A. (2006). Reasons without demands: Rethinking rightness. In J. Dreier (Ed.), *Contemporary debates in moral theory* (pp. 38–54). Oxford: Blackwell.
- Nozick, R. (1974). *Anarchy, state, and utopia*. Oxford: Blackwell.
- Peterson, M. (2013). *The dimensions of consequentialism: Ethics, equality and risk*. Cambridge: Cambridge University Press.
- Powell, B., & Zwolinski, M. (2012). The ethical and economic case against sweatshop labor: A critical assessment. *Journal of Business Ethics*, 107(4), 449–472.
- Robinson, J. (1962). *Economic philosophy*. Bungay: Penguin.

- Roemer, J. (1982). *A general theory of exploitation and class*. Cambridge, MA: Harvard University Press.
- Ross, R. (2004). *Slaves to fashion: Poverty and abuse in the new sweatshops*. Ann Arbor, MI: University of Michigan Press.
- Sample, R. (2003). *Exploitation: What it is and why it's wrong*. Boulder, CO: Rowman and Littlefield.
- Scheffler, S. (1988). *Consequentialism and its critics*. Oxford: Oxford University Press.
- Sobel, D. (2007). The impotence of the demandingness objection. *Philosopher's Imprint*, 7(8), 1–17.
- Steiner, H. (1984). A liberal theory of exploitation. *Ethics*, 94(2), 225–241.
- Steiner, H. (1987). Exploitation: A liberal theory amended, defended and extended. In A. Reeve (Ed.), *Modern theories of exploitation* (pp. 132–148). London: Sage.
- Temkin, L. (2012). *Rethinking the good: Moral ideals and the nature of practical reasoning*. Oxford: Oxford University Press.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458.
- Vallentyne, P. (1993). The connection between prudential and moral goodness. *Journal of Social Philosophy*, 24(2), 105–128.
- Veneziani, R. (2013). Exploitation, inequality, and power. *Journal of Theoretical Politics*, 25(4), 526–545.
- Voorhoeve, A. (2013). Vaulting intuition: Temkin's critique of transitivity. *Economics and Philosophy*, 29(3), 409–423.
- Vrousalis, N. (2013). Exploitation, vulnerability, and social domination. *Philosophy and Public Affairs*, 41(2), 131–157.
- Wertheimer, A. (1996). *Exploitation*. Princeton: Princeton University Press.
- Wertheimer, A. (2007). Review: Ruth J. Sample, "Exploitation: What it is and why it's wrong". *Utilitas*, 19(2), 259–261.
- Zwolinski, M. (2007). Sweatshops, choice, and exploitation. *Business Ethics Quarterly*, 17(4), 689–727.